first frame of the extracted feature data for the candidate content. The feature data indexed by the reference index and candidate index are compared **304**. A test **306** is performed to determine whether the comparison indicates they match. The term "match," as used herein, includes exact matches between extracted reference and candidate feature data, as well as matches within a desired tolerance threshold. For example, one may try to trick an authentication system by degrading the candidate content before trying to share it. This may result in a degradation of extracted feature data for the candidate content. However, the match threshold may be set such that degraded feature data nonetheless matches the reference feature data. (See FIG. 4.)

If the test **306** does not indicate a match, another test **308** is performed to determine whether the reference index points to the last frame of the feature data for the reference content. If yes, then there has not been a match between the extracted feature data for the reference content and the candidate content, and processing ends **310**. If the test **308** indicates more frames, then the reference index is incremented **312** to the next frame, thus moving the sliding-window for continuing comparison **304**.

If the test **306** indicates a match between the extracted feature data, a test **314** is performed to determine whether the last frame of the candidate feature data has been reached. If yes, then there has been a match between the extracted feature data for the reference content and the candidate content, and processing ends **314**. As discussed above, various responses may be taken on a successful match, including denying the sharing, providing purchase information for the sharing, or other action. If the test **314** indicates the last frame has not been reached, then the candidate index is

incremented **316** to the next frame, thus continuing comparison **304** within the current position of the current sliding-window on the reference feature data.

Although the illustrated embodiment shows a single-threaded sliding-window comparison, it will be appreciated that the comparison may be parallelized for a multiprocessor computing device. For example, if there are N reference content entries in the database, then N processors could be assigned to compare candidate against all reference content in parallel.

Alternatively, assuming there are as many processors as there are frames in the reference content, and the extraction process employed allows matching a potential starting frame (there may be more than one apparent starting frame) against extracted candidate content feature data, then identifying the correct starting frame can be performed in a single operation.

Similarly, assuming sufficient processors, and the extraction process employed allows matching a potential starting frame against extracted candidate content feature data, a comparison between the candidate and reference content feature data can be performed in a single operation, plus a test to see if any comparisons were positive. For example, If the reference content is M frames, and the extracted candidate content is N frames, then N processors can be assigned to the first M-N reference frames, where $(M\text{-}N)$xN comparisons are performed simultaneously. A test may be performed to see if any of the comparisons indicated a match. Of course, other parallel testing configurations may be performed with fewer processors.

FIG. 4 illustrates a comparison between 11 reference contents against themselves, 3 degraded copies of the reference content, and 10 entirely different content, where match confidence values are plotted on both the X and Y axes of the chart **400**. Illustrated is a scatter plot of match scores ranging between 0 and 1 plotted against themselves. Thus both the X-axis **402** and Y-axis **402** of the chart illustrate match confidence values for the degree to which extracted feature data for particular candidate content matches extracted feature data for the reference content. A score at the origin (0,0) **406** is a perfect match, representing comparing content against itself. Scores moving away from the origin represent matches having less match confidence. A score at (1,1) **416** indicates a total mismatch.

In the illustrated embodiment, reference content was intentionally degraded by using lossy video compression at different quality levels; using a high level of compression resulted in the greatest degradation. Video compression is typically used for storing and sharing video, and therefore is appropriate for identifying attempts to engage in illicit sharing of reference content. It will be appreciated, however, that many other alterations may be made to distort the reference content, and these different alteration techniques may be accounted for when matching reference and candidate content. For example, one may re-encode the reference content with a poor encoding engine. In this latter example, since candidate content needs to be decoded in order to extract feature data, a poor encoding technique results in degraded output, display window reduced size, poorer audio quality.

The chart clusters **408-412** illustrate three comparisons between the 11 reference content videos and three different degradations of each of reference content.